

# G.718 i kompresija govornog signala u telekomunikacijskim sustavima

---

**Kulešević, Franko**

**Undergraduate thesis / Završni rad**

**2021**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:200:411139>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-12-27**

*Repository / Repozitorij:*

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU**

**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I  
INFORMACIJSKIH TEHNOLOGIJA**

**Stručni studij**

**G.718 I KOMPRESIJA GOVORNOG SIGNALA**

**Završni rad**

**Franko Kulešević**

**Osijek, 2021.**

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA  
I INFORMACIJSKIH TEHNOLOGIJA OSIJEK**Obrazac Z1 - Obrazac za ocjenu završnog rada**

Osijek, 2021

**Odboru za završne i diplomske ispite****Prijedlog ocjene završnog rada**

Ime i prezime studenta:

Smjer:

Mat. br. studenta, godina upisa:

Mentor:

Sumentor:

Naslov završnog rada:

Primarna znanstvena grana rada:

Sekundarna znanstvena grana (ili polje) rada:

Predložena ocjena završnog rada:

Kratko obrazloženje predložene ocjene:

Potpis sumentora:

Potpis mentora:

Dostaviti:

1. Studentska služba

**Korekcija ocjene Odbora za završne i diplomske ispite**

Odbor za završne i diplomske ispite Fakulteta elektrotehnike, računarstva i informacijskih tehnologija Osijek, temeljem članka 11. Pravilnika o završnim ispitima na preddiplomskim i stručnim ispitima Fakulteta elektrotehnike, računarstva i informacijskih tehnologija Osijek donio je odluku o korekciji konačne ocjene završnog rada na \_\_\_\_  
sjednici Odbora održanoj dana \_\_\_\_\_ godine.

Korigirana konačna ocjena završnog rada: \_\_\_\_\_

Potpis predsjednika Odbora:  
\_\_\_\_\_

Dostaviti:

1. Studentska služba



**FERIT**

FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA  
I INFORMACIJSKIH TEHNOLOGIJA OSIJEK

## IZJAVA O ORIGINALNOSTI RADA

Osijek,

Ime i prezime studenta:

Studij :

Mat. br. studenta, godina upisa:

Ovom izjavom izjavljujem da je rad pod nazivom:

izrađen pod vodstvom mentora

i sumentora

moj vlastiti rad i prema mom najboljem znanju ne sadrži prethodno objavljene ili neobjavljene pisane materijale drugih osoba, osim onih koji su izričito priznati navođenjem literature i drugih izvora informacija.

Izjavljujem da je intelektualni sadržaj navedenog rada proizvod mog vlastitog rada, osim u onom dijelu za koji mi je bila potrebna pomoć mentora, sumentora i drugih osoba, a što je izričito navedeno u radu.

Potpis studenta:

# SADRŽAJ

1. UVOD .....	1
1.1 ZADATAK ZAVRŠNOG RADA .....	2
2. OSNOVNE KARAKTERISTIKE .....	3
3. PRINCIP RADA .....	5
3.2. Koder .....	6
3.2.1 Princip rada koder .....	7
3.2.2 CELP algoritam i analiza slojeva u G.718 preporuci .....	9
3.3 Dekoder .....	11
3.3.1 Princip rada dekodera .....	12
4. PRIMJENA I PREDNOSTI .....	14
5. USPOREDBA S OSTALIM PREPORUKAMA .....	16
6. ZAKLJUČAK .....	19
LITERATURA .....	20
SAŽETAK .....	22
ABSTRACT .....	23
ŽIVOTOPIS .....	24

# 1.UVOD

U tradicionalnim telekomunikacijskim mrežama kompresija govornog signala se vrši po ITU preporuci G.711. G.711 je standard kodiranja govornog signala dizajniran za potrebe telefonske infrastrukture, a osnovna karakteristika mu je brzina prijenosa podataka od 64 kbit/s [1]. Taj standard obuhvaća govorne signale unutar spektra od 300-3400 Hz te ih uzorkuje brzinom od 8000 uzoraka u sekundi. G.711 se službeno zove pulsno-kodna modulacija zvučnih frekvencija (eng. *Pulse code modulation of voice frequencies*) [1]. Pulsno-kodna modulacija (u daljnjem tekstu PCM) je proces u kojem se analogni signal pretvara u digitalni signal gdje je svaki diskretni član digitalnog signala predstavljen kodnom riječi. Amplituda analognog signala se uzorkuje konstantnom brzinom pri uniformnim intervalima, a svaki uzorak je kvantiziran na način da se uzima najbliža vrijednost trenutnoj analognoj vrijednosti iz koje se uzima uzorak. Međutim, zbog zahtjeva na preporukama na uštedi spektra pojavila se potreba za novim metodama kompresije. Jedna od tih metoda je preporuka G.718. Preporuka u sebi opisuje funkcije kodiranja te dekodiranja nekog signala koji može biti vizualne ili zvučne prirode te funkcije kompresije signala. Funkcije kodiranja predstavljaju pretvorbu neke informacije u podatkovni oblik koji je prikladan za prijenos kroz komunikacijski medij, dok funkcija dekodiranja predstavlja pretvorbu te informacije iz podatkovnog oblika u njen izvorni oblik prije procesa kodiranja. U tom procesu pretvorbe informacije iz jednog oblika u drugi, može se raditi proces kompresije informacije u podatkovnom obliku (u daljnjem tekstu podatak). Kompresija ili sažimanje je proces kodiranja informacije na način da se koristi manje bitova podataka od broja bitova koje sadrži njen izvor [2]. S time se postiže smanjenje širine pojasa, smanjenje prostora pohrane tih podataka te smanjeno vrijeme prijenosa kroz neki komunikacijski kanal. Kodeci se dijele na kodeke s gubicima u kompresiji (eng. *lossy*) te na kodeke bez gubitaka u kompresiji (eng. *lossless*) [2]. Kodeci s gubicima u kompresiji se koriste u većem broju, poglavito zbog njihovog svojstva da mogu očuvati kvalitetu izlaznog signala na dovoljno visokoj razini, gdje se uklanjaju podaci koji se smatraju redundantnim prilikom procesa kompresije. Na taj način se smanjuje širina pojasa prilikom kodiranja i povećavaju performanse samog kodeka. No posljedično se smanjuje kvaliteta usluge kako bi se maksimalno komprimirali ti podaci. Kodeci bez gubitaka u kompresiji se koriste za spremanje podataka u komprimiranom obliku, a na način da se istovremeno zadrže sve informacije koje su bile prisutne u izvornom informacijskom toku. Kodek G.718 je kodek s gubitcima tijekom kompresije. G.718 je preporuka od organizacije ITU-

T koja opisuje uskopojasni i širokopojasni ugradbeni promjenjivi algoritam kodiranja govora i audio signala koji ima brzine kodiranja od 8 do 32 kbit/s te koji je dizajniran s namjerom da bude robustan na gubitke okvira (eng. *frames*). S time se postiže povećanje kvalitete govora kada se koristi u sprezi s IP mrežnim protokolom na fiksnim, bežičnim i mobilnim mrežama. U ovom radu će se prvotno opisivati proces koji je doveo do razvitka ove preporuke u obliku kojeg imamo danas. U drugom dijelu rada će se opisivati karakteristike kodeka ili preporuke te će se pristupiti detaljnijoj obradi njegovih podsustava. U nastavku će se opisivati princip rada te primjena kodeka, a na kraju će se dati okvirni pregled usporedbe s ostalim standardiziranim audio kodecima.

## **1.1 ZADATAK ZAVRŠNOG RADA**

Poznato je da se u tradicionalnim telekomunikacijskim mrežama kompresija govornog signala vrši po ITU preporuci G.711. U modernijim sustavima koji prenose govorni signal došlo je, zbog uštede na spektru do pojave novih metoda kompresije. Te nove metode opisane su u ITU-T preporukama. Jedna od njih je G.718. Zadatak je proučiti povijest i razloge kompresije govornog signala te detaljno obraditi CELP algoritam i način kompresije definiran u preporuci G.718.

## 2. OSNOVNE KARAKTERISTIKE

Razvoj G.718 kodeka je započeo 1999. godine pokušajem standardizacije procesa kodiranja s promjenjivom brzinom prijenosa podataka [3]. Dvije paradigme su se uzele u obzir, a to su promjenjiva brzina prijenosa podataka (eng. *variable bit rate*) te ugradbena brzina prijenosa podataka (eng. *embedded bit rate*). Preporuka ili kodek sadrži u sebi tzv. slojeve. Sloj predstavlja razinu implementacije na softverskoj, odnosno hardverskoj razini koja dostavlja govorni signal na izlaz kodeka u skladu s parametrima definiranim u preporuci. Jedan od njih je osnovni ili jezgri sloj. Paradigma konstantne brzine prijenosa podataka opisuje jezgri sloj s konstantnom brzinom prijenosa podataka za signale koji u sebi sadrže gubitke informacija uslijed pauza koje nastaju tokom govora osobe. Povijesno gledano je paradigma promjenjive brzine prijenosa podataka bila prikladna za CDMA sustave (eng. *Code Division Multiple Access*) dok se u drugu ruku, paradigma s ugradbenom brzinom prijenosa podataka već koristila unutar ITU-T preporuka. Potonja sadrži jezgri sloj koji je predviđen da radi na frekvenciji uzorkovanja od 8 kHz koja se inače koristila u staroj tehnologiji telefonije. Viši slojevi sadrže implementaciju koja omogućuje dogovorenu brzinu prijenosa informacija te posjeduju skalabilnost ili nadogradivost širine pojasa. Kako je razvoj sazrijevao, pojavilo se uvriježeno mišljenje da preporuke temeljene na jezgrama koje obrađuju širokopolasni ili uskopolasni signal se ne trebaju koristiti. Također je zavladao mišljenje da rješenja temeljena na suradnji s AMR-WB standardom kodiranja, zatim s G.722.2 standardom kodiranja te konačno s VMR-WB jedina kao takva imaju smisla za implementaciju sljedeće generacije kodeka. Neke su organizacije izrazile zabrinutost u tom vremenu da će takvo rješenje biti novi kodek koji radi brzinom kodiranja od oko 8 kb/s s uskopolasnim signalima. Budući da im to nije odgovaralo zbog zahtjeva tržišta, podiglo se pitanje početkom 2004. godine oko promjene osnovnih zahtjeva implementacije novog kodeka. Taj kodek bi ipak radio kodiranje signala spomenutom brzinom uz uvjet da se kao kompromis ostavi mogućnost korištenja jezgre koja radi s uskopolasnim signalom. S vremenom kako je tekao razvoj dodao se novi zahtjev za povećanje tolerancije na smetnje u komunikacijskom kanalu. Nakon daljnjeg razvoja implementacije kodeka, određen je termin početne faze evaluacije implementacije kodeka u ožujku 2007. godine kako bi se odabrala početna točka nakon koje bi se započela daljnja optimizacija. U tom periodu pojavila su se četiri kodeka koji su bili kandidati za evaluaciju u fazi selekcije u tom periodu. Rješenje koje su zajednički razvile tvrtke Ericsson, Motorola, Nokia, Texas Instruments te VoiceAge su pobijedile u natjecanju te je njihovo rješenje izabrano kao početna točka za daljnju optimizaciju.



Pokušaj standardizacije u ovom periodu je drugačiji u odnosu na pokušaje u prošlim periodima na način da se ovaj put dopustila optimizacija svim tvrtkama koje su bile zainteresirane za poboljšanje osnovne verzije kodeka. Devet drugih tvrtki su izrazile namjeru za sudjelovanje u razvoju kodeka, no samo su tehnologije koje su razvijene od strane firmi Huawei, France Telecom te Qualcomm uključene u konačno rješenje. Faza optimizacije je završena u travnju 2008. godine. To rješenje je u konačnici prihvaćeno u lipnju te godine kao ITU-T preporuka G.718. Odgovarajući opis plutajućeg zareza (eng. *floating point description*) je odobren u ožujku 2009. godine kao dodatak A preporuci G.718 (implementacija je pisana u C programskom jeziku). U posljednjoj fazi standardizacije preporuke G.718 karakterizacija kodeka je završena procjenom ponašanja implementacije u raznim uvjetima, posebice testiranjem pomoću korištenja komunikacijskog kanala koji je bio zasićen šumom. Određeni rezultati testiranja se nisu uzeli u obzir budući da tada nisu postojali jasni ciljevi kvalitete usluga koje trebaju biti isporučene u takvim uvjetima (u ovom slučaju reproducirani govorni signal). Tokom razvoja G.718 preporuke, vidjelo se da kodek može uzorkovati signale zvuka s monofonim signalima koji su ograničeni na frekvencijski pojas u iznosu širine 7 kHz. Stoga, u paraleli s fazom optimizacije G.718, standardizacija novog dodatka preporuci je objavljena, a koja može omogućiti kodeku da reproducira zvučne signale u frekvencijskim pojasima u iznosu od 50 Hz do 14000 Hz. Taj dodatak se razvija istovremeno za preporuke G.718 te G.729.1. Za G.718 dodatak preporuci će raditi na brzinama kodiranja podataka u iznosu od 28 kb/s do 64 kb/s ovisno o širini pojasa te broju ulaznih kanala. Očekivani završetak standardizacije dodatka je planiran 2010. godine.

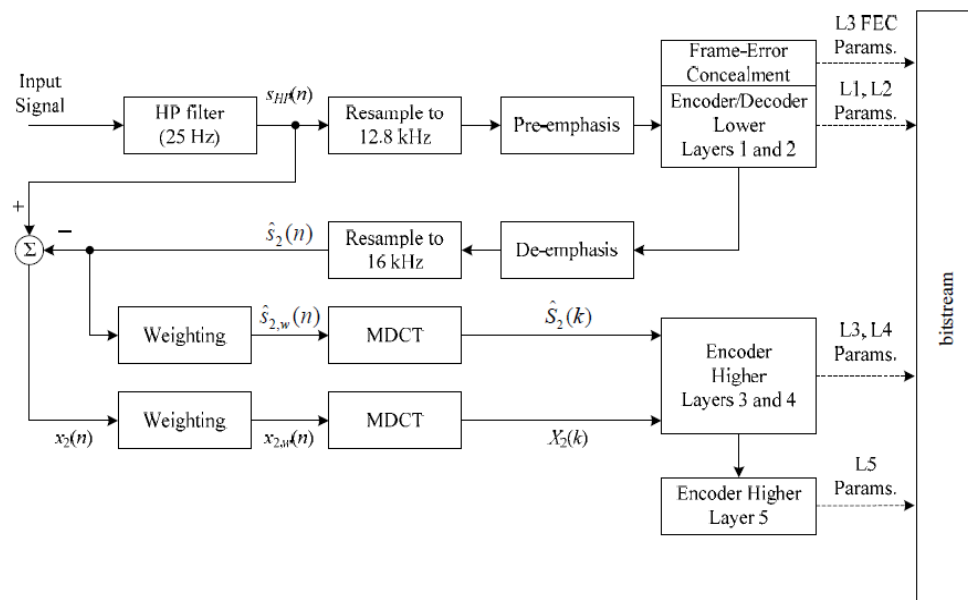
### 3. PRINCIP RADA

Preporuka ITU-T G.718 opisuje uskopojasni (eng. *narrowband*) i širokopojasni (eng. *wideband*) algoritam kodiranja zvučnog signala (npr. ljudski govor) pri brzinama kodiranja u intervalu od 8 do 32 kbit/s [4]. Preporuka je dizajnirana da bude otporna na gubitke u prijenosu ili tijekom same obrade podataka. Ovaj kodek omogućuje visoku kvalitetu govora u niskopropusnom frekvencijskom području pri nižim brzinama prijenosa podataka. Također to vrijedi i za visoku kvalitetu govora u širokopojasnom frekvencijskom području pri svim dostupnim brzinama prijenosa podataka. G.718 kodek je dizajniran da bude robustan na smetnje. Robusnost se manifestira filtriranjem šuma u određenim slojevima, a s time se postiže poboljšanje kvalitete govora. Preporuka se najviše koristi u aplikacijama koje koriste IP protokol u fiksnoj, bežičnoj i mobilnoj mrežnoj infrastrukturi. Posjeduje ugrađenu skalabilnu strukturu, koja omogućava maksimalnu fleksibilnost u transportu IP paketa koji sadrže zvučne podatke. Dodatno je omogućena fleksibilnost oko promjene brzine prijenosa podataka na način da strana dekodera može odbaciti pakete po potrebi. Strana koda u implementacijskom smislu sadrži pet slojeva koji obrađuju podatke i vrše funkcije transformacije nad podacima određenim brzinama kako je definirano u samoj preporuci. Te brzine su: 8, 12, 16, 24 te 34 kbit/s. Koder može primiti na ulazu širokopojasne uzorkovane signale frekvencijom uzorkovanja od 16 kHz, te niskopojasne signale uzorkovane na frekvencijama u iznosu od 16 ili 8 kHz. Dekoder reproducira izlazni signal frekvencije u iznosu od 16 kHz. Ulazni signal se uzorkuje frekvencijom od 16 kHz, a širina pojasa je na intervalima od 300-3400 Hz pri brzini od 8 do 12 kbit/s te 50-7000 Hz pri brzini od 8 do 32 kbit/s. Jezgri sloj kodeka omogućava onu kvalitetu govora koja je ekvivalentna kvaliteti govora ITU-T preporuke G.722.2 kodeka pri 12.65 kbit/s u širokopojasnom području [4]. U niskopojasnom području jezgri sloj kodeka omogućava kvalitetu govora ekvivalentnu kvaliteti govora ITU-T preporuke G.729 pri brzini od 11.8 kbit/s. G718 kodek radi na okvirima trajnosti 20 ms i ima maksimalno kašnjenje definirano algoritmom u iznosu od 42.785 ms za ulazne i izlazne širokopojasne signale. Maksimalno kašnjenje definirano algoritmom za niskopojasne ulazne i izlazne signale je 43.785 ms. Kodek može raditi u načinu rada s niskim iznosom kašnjenja kada su maksimalne brzine prijenosa podataka koda i dekoda podešene na brzinu od 12 kbit/s. U tom slučaju maksimalno kašnjenje definirano algoritmom je skraćeno na 10 ms. Algoritam koji se koristi u ovom kodeku se temelji na strukturi kodiranja koja je podijeljena u dvije faze, a to su donja i gornja faza. Donja faza se temelji na dva sloja koji koriste CELP (eng. *code-excited linear prediction*) za pojase iznosa od

50 do 6400 Hz gdje jezgri sloj radi optimizaciju kodiranja na temelju klasifikacije ulaznog signala za svaki okvir koji sadrži podatke ulaznog signala. Gornja faza kodira signal iz nižih slojeva koristeći MDCT (eng. *modified discrete cosine transformation*) transformaciju. Kodek se sastoji od više elemenata, a to su koder i dekodek koji su razdvojeni komunikacijskim kanalom. Blok dijagrami su dani u njihovim narednim potpoglavljima. Također je opisana i funkcionalnost članova prikazanih u blokovima koji će se također detaljnije obraditi.

### 3.2. Koder

Koder je entitet koji pretvara signal u kodiranu poruku, odnosno u kod koji je prikladan za prijenos kroz komunikacijski kanal. Kod je sustav pravila koji se koristi za pretvorbu jednog signala (u ovom slučaju govorni signal) u drugi oblik signala koji je prikladan za daljnju obradu [5]. Primjeri daljnje obrade su prijenos kroz komunikacijski kanal ili pohrana podataka u neki medij. Proces koji označava pretvorbu jednog oblika signala u drugi se naziva kodiranje. Dekodiranje je proces inverzan kodiranju, odnosno ulazni signal je kodirani signal, a izlazni signal je originalni govorni signal koji se reproducira na prijemnom kraju. Na slici 1 je prikazan shematski blok dijagram kodera zajedno s nazivima pojedinih članova tog dijagrama [4].



Slika 1: Strukturni blok dijagram kodera

Slika 1 prikazuje strukturni blok dijagram kodera s različitim slojevima. Slojevi su označeni na sljedeći način: L1 predstavlja jezgri sloj, dok četiri preostala sloja (L2-L5) predstavljaju dodatne slojeve [6]. Slojevi se pored implementacije funkcionalnosti mogu promatrati i kroz njihov izlaz, a to je govorni signal, odnosno njegova kvaliteta. Kvaliteta govornog signala se može raspodijeliti na sljedeći način. Jezgri sloj predstavlja osnovnu minimalno dogovorenu kvalitetu govora po preporuci, dok ostali slojevi proizvode kvalitetniji signal, zavisno od razine koja se promatra. Brzine prijenosa podataka za jezgri sloj iznose 8 kbit/s dok ako se uključe i ostali slojevi dostižu se brzine do 32 kbit/s. Iz slike se može vidjeti da se donja dva sloja (L3 te L4 i L5) primjenjuju u signalu čiji su harmonici dodatno pojačani kako bi se lakše uklonio šum [4]. Zatim su ponovno uzorkovani pri frekvenciji uzorkovanja od 12.8 kHz, dok se gornja tri (L1, L2 i L3 FEC) sloja uzorkuju frekvencijom uzorkovanja u iznosu od 16 kHz. Za prijenos signala komunikacijskim kanalom je dovoljan isključivo jezgri sloj L1 koji je dovoljan i na strani dekodera za uspješno dekodiranje zvučnog sadržaja [4]. Ostali slojevi služe samo za pojačavanje kvalitete zvuka. Stoga, jezgri sloj osigurava osnovnu komunikaciju dok ostali slojevi služe za poboljšanje percepcije kvalitete zvuka. Dva su slučaja ulaznog signala, a to su uskopojasni ili širokopojasni ulazni signal. Jezgri sloj (L1) koristi CELP tehnologiju gdje se govorni signal modelira pobudnim signalom koji je prošao LP (eng. *linear prediction*) filter. Analiza vršnih vrijednosti (eng. *open-loop pitch analysis*) se provodi s algoritmom koji traži vrijednosti harmonika koji pripadaju govoru i uzimaju vrijednosti koji se nalaze ispred i iza tih vršnih vrijednosti. To se radi u svrhu kasnijeg određivanja izlaznog govornog signala na strani dekodera. U sljedećem potpoglavlju su zasebno objašnjeni elementi unutar bloka dijagrama kodera te sam princip rada kodera i kompresija govornog signala.

### 3.2.1 Princip rada kodera

Visokopropusni filter predstavlja prvi element kodera u koji ulazi širokopojasni ili uskopojasni zvučni signal [4]. Ulazni signal je prethodno uzorkovan s frekvencijom od 8 ili 16 kHz, ovisno o danoj pojasnoj širini ulaznog signala. Ovaj filter ima funkciju uklanjanja harmonika koji pripadaju nižim frekvencijama koje su neželjene za kasnije kodiranje. Zatim signal dolazi u blok za ponovno uzorkovanje (eng. *resampling*) koji ima funkciju ponovnog uzorkovanja novom frekvencijom od 12.8 kHz [4]. S time se postiže manja širina pojasa zvučnog signala potrebna za slanje komunikacijskim kanalom. Zatim takav signal ulazi u blok za pojasno propusno filtriranje (eng. *pre-emphasis*) koji se koristi za filtriranje određenog pojasa frekvencijskih harmonika ulaznog signala. Filter je opisan prijenosnom funkcijom [4]. Nadalje,

odrađuje se spektralna analiza unutar kodera koja se koristi za detekciju pojave govornog signala te smanjenje šuma [4]. Za spektralnu analizu se koristi diskretna Fourierova transformacija. Analiza frekvencija se radi dva puta po okviru koristeći brzu Fourierovu transformaciju (u daljnjem tekstu FFT). Početak prvog prozora se stavlja na početku trenutnog okvira. Sljedeći prozor se postavlja na sredinu okvira. Kompleksni matematički proračun se koristi za obradu, ali rezultat FFT analize je rezultatni frekvencijski spektar koji je podijeljen na kritične pojaseve [4]. Dobiva se dvadeset frekvencijskih pojaseva u frekvencijskom intervalu od 0 – 6400 Hz. Prema tim pojasevima se kasnije određuju parametri energije okvira, detekcija uskopojasnih ulaznih pojaseva, detekcija aktivnosti signala itd. G.718 kodek je opremljen s integriranim algoritmom smanjenja šuma koji radi s ulaznim signalom koji je izlaz iz bloka za ponovno uzorkovanje [4]. Takav signal se koristi za procjenu nekih parametara kodera, ali za kodiranje se koristi originalni ulazni signal. Parametri koji se mogu estimirati iz signala filtriranog od šuma su oni koji se koriste kao baza za odabir moda kodiranja. Iznimka je kad je operacija kodera ograničena da radi sa slojem L2 uključivo s jezgrenom slojem. Analizom spektra pomoću metode linearne predikcije se određuju koeficijenti sinteznog filtra CELP modela koji je detaljno obrađen u narednom potpoglavlju [4]. Ti koeficijenti se kasnije koriste za određivanje linearno-predikcijskih koeficijenata pomoću Levinson - Durbin algoritma. Zatim se ti koeficijenti transformiraju u ISP parametre (eng. *immittance spectral pairs*) te posljedično u ISF parametre (eng. *immittance spectral frequencies*) za kasniju kvantizaciju i interpolaciju. Interpolirani kvantizirani i nekvantizirani koeficijenti se pretvaraju kasnije nazad u linearno predikcijske koeficijente kako bi se kasnije konstruirali težinski filteri za svaki podokvir. Što se tiče jezgrenog sloja, da bi se postigle maksimalne performanse kodiranja pri brzini od 8 kbit/s, jezgreni sloj koristi algoritam klasifikacije signala te četiri različita moda kodiranja koji su jedinstveni za svaku klasu signala [4]. To su bezvučni kodirajući mod (eng. *unvoiced coding mode*), zvučni kodirajući mod (eng. *voiced coding mode*), prijelazni kodirajući mod (eng. *transition coding mode*) te generički kodirajući mod (eng. *generic coding mode*). Algoritam klasifikacije signala koristi mnoštvo parametara, neki od njih su optimizirani odvojeno za širokopojasne i uskopojasne ulazne signale. S ovim modovima rada postiže se maksimalna efikasnost kodiranja zvučnog signala gdje se primjenjuju različiti matematički proračuni ovisno o klasifikaciji signala koji se trenutno analizira za proces kodiranja. Nadalje, od viših slojeva, postoji ACELP sloj koji se koristi za poboljšanje izlaza iz jezgrenog sloja [4]. ACELP sloj koristi ACELP (eng. *algebraic code excited linear prediction*) algoritam kodiranja. Taj algoritam je u suštini LPC algoritam koji se temelji na CELP metodi i ima algebarsku strukturu. Jedna od glavnih značajki ACELP-a je da posjeduje veliki rječnik koji je veći od 50 bita s čime je omogućena obrada veće količine

informacija govornog signala. I ovaj algoritam će se detaljnije obraditi u sljedećem potpoglavlju. Sljedeći sloj, za analizu signala u uvjetima gdje se događa potencijalno brisanje okvira te koristi se za očuvanje kvalitete u uvjetima kada se događaju gubici okvira uslijed smetnji u komunikacijskom kanalu [4]. Ovaj sloj u podatkovni tok ubacuje informacije relevantne za dekođer, a to su informacije o parametrima koji opisuju brisanje okvira. S tim se postiže neometani rad dekođera uslijed otežanih uvjeta unutar komunikacijskog kanala. Viši slojevi općenito nastupaju nakon što završi obrada signala u nižim slojevima. Razlikovni signal iz druge faze CELP kodiranja unutar sloja L2 se dalje kvantizira u slojevima L3, L4 i L5 pomoću MDCT (eng. *modified discrete cosine transform*).

### 3.2.2 CELP algoritam i analiza slojeva u G.718 preporuci

U ovom poglavlju će se opisati CELP algoritam i ostali algoritmi koji se koriste za kodiranje govora s obzirom na sloj kodeka. U sljedećoj tablici se može naći popis algoritama koji se koristi u ovisnosti o sloju koji se koristi prilikom obrade ulaznog signala. Tablica se temelji na podacima dostupnim u literaturi [4] i [7].

Sloj	Brzina prijenosa	Algoritam	Frekvencija uzorkovanja
L1	8 kb/s	CELP	12.8 kHz
L2	+ 4 kb/s	ACELP	12.8 kHz
L3	+ 4 kb/s	FEC/MDCT	12.8 kHz ili 16 kHz
L4	+ 8 kb/s	MDCT	16 kHz
L5	+ 8 kb/s	MDCT	16 kHz

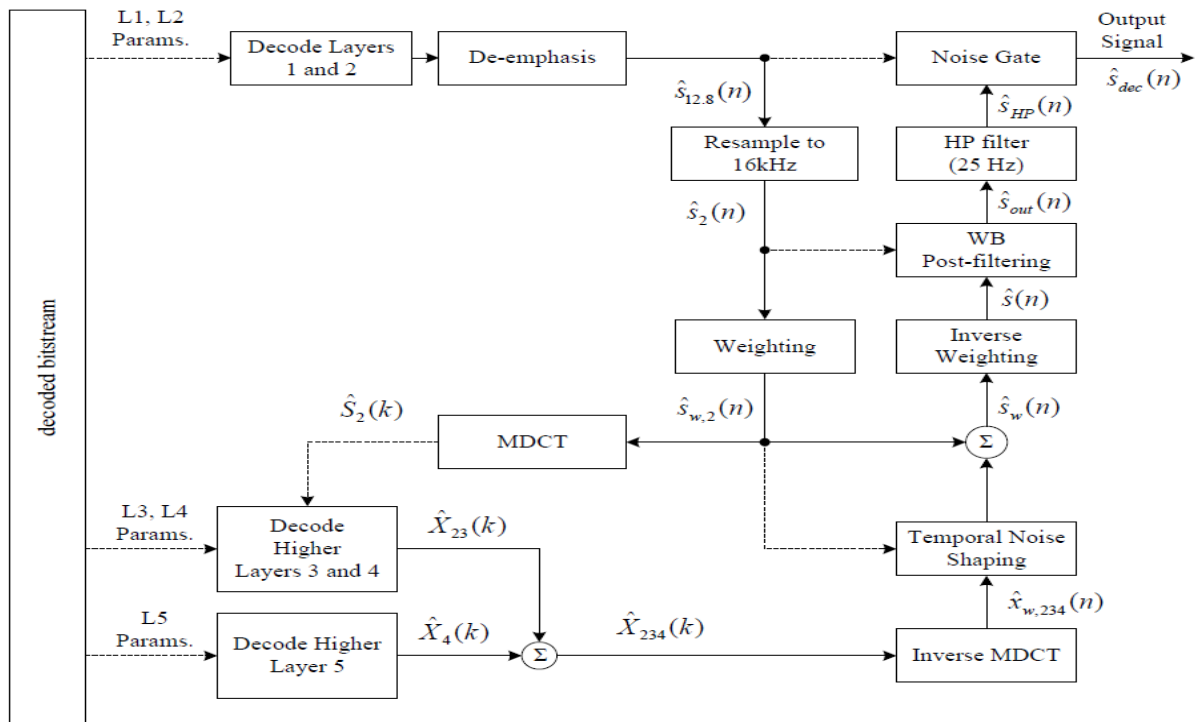
**Tablica 1:** Prikaz slojevite strukture G.718 kodeka

Prema tablici 1 se može vidjeti koji algoritam kodiranja se koristi, koje frekvencije uzorkovanja signala te koje brzine prijenosa podataka s obzirom na sloj. Sloj L1 koristi CELP algoritam kodiranja pri frekvenciji uzorkovanja od 12.8 kHz. Ovaj algoritam se pokazao efikasnim u kodiranju govornog signala pri nižim brzinama prijenosa podataka [7]. CELP algoritam je linearni prediktivni algoritam kodiranja predstavljen 1985. godine [8]. Ima bolje performanse u odnosu na konkurentne algoritme, poput RELP algoritma (eng. *residual-excited linear prediction*), LPC algoritma (eng. *linear predictive coding*). Zajedno sa svojim izvedenim algoritmima poput ACELP algoritma koji se koristi u G.718 preporuci, trenutno je najčešće

korišten algoritam kodiranja govora. CELP algoritam se temelji na nekoliko ideja [8]. Te ideje su korištenje modela izvor-filter (eng. *source-filter model*) kroz linearnu predikciju, zatim korištenje adaptivnog te fiksnog rječnika kao ulaz za LP (linearno-predikcijski) model, pretraga težinskih faktora percepcije govornog signala te primjena vektorske kvantizacije. CELP se koristi i na strani kodera i na strani dekodera, a na strani dekodera će se detaljnije objasniti u narednim potpoglavljima. Glavni princip na kojem se temelji CELP je analiza na temelju sinteze (eng. *Analysis-by-Synthesis*). Glavna značajka tog principa je da se proces kodiranja izvodi na način da se pomoću percepcije optimizira dekodirani signal (predstavlja sintezu) u zatvorenoj petlji [8]. Pod percepcijom se podrazumijeva ljudska percepcija govornog signala. Idealni reproducirani govorni signal na izlazu iz dekodera bi bio onaj koji bi izabrali slušatelji kada bi preslušali sve moguće kombinacije bitova i odabrali onu kombinaciju za koju bi se složili da je idealna, no to nije moguće zbog velikog broja kombinacija. Kako bi se ostvarilo kodiranje u realnom vremenu u uvjetima ograničenih računalnih resursa prilikom obrade signala, CELP pretraga se dijeli na nekoliko sekvencijalnih pretraga koristeći jednostavne težinske funkcije. Tipično se kodiranje provodi u sljedećem redu: prvotno se računaju linearni predikcijski koeficijenti te se kvantiziraju, zatim se provodi pretraga rječnika kako bi se odabrale sekvence bitova za slanje komunikacijskim kanalom. Većina modernih kodeka oblikuju šum tako da se pojavi samo u onim frekvencijskim područjima gdje ga ljudsko uho ne može detektirati. Sloj L2 koristi ACELP algoritam za kodiranje govornog signala. To je LPC algoritam koji se temelji na CELP algoritmu te posjeduje algebarsku strukturu rječnika [9]. Glavna značajka rječnika je njegova veličina, a može biti veći od 50 bitova. Kako slojevi L1 i L2 služe za kodiranje govornog signala stoga imaju smanjenu efikasnost za glazbene signale [7]. G.718 preporuka podrazumijeva kodiranje generičkog zvuka, koji u sebi sadrži i govorni signal, te stoga za kodiranje zvuka čiji frekvencijski spektar ne pripada govornom zvuku (signalu) primjenjuje više slojeve (L3, L4, L5). Viši slojevi prema tablici 1 koriste MDCT osim sloja L3 koji sadrži u sebi i FEC (eng. *frame erasure concealment*). FEC mehanizam predstavlja mehanizam skrivanja okvira koji se koristi za prilagodbu parametara rječnika. Rječnik u sebi sadrži bitove koji se mijenjaju s obzirom na podatke koji stižu unutar okvira, a ti okviri (eng. *frames*) predstavljaju podatkovni tok koji traje 20 ms. Ponekad se događaju gubici okvira u otežanim uvjetima gdje je prisutan visok šum, pa taj mehanizam nastupa u tom slučaju i prilagođava sadržaj rječnika za efikasnije kodiranje zvučnog signala. MDCT je modificirana diskretna kosinusna transformacija koja se koristi kod zvukova koji pripadaju van govornog područja (npr. glazba). MDCT se temelji na preklapanju podataka.

### 3.3 Dekoder

Na slici 2 je prikazan shematski blok dijagram dekodera zajedno s nazivima pojedinih članova tog dijagrama [4].



**Slika 2:** Strukturni blok dijagram dekodera

Slika 2 prikazuje strukturni blok dijagram dekodera [4]. Dekoder iz podatkovnog toka raznim mehanizmima koji će biti opisani u ovom potpoglavlju i u narednim potpoglavljima generira izlaz koji će biti govorni signal. U svakom okviru (koji traje 20 ms) nalaze se podaci koji su generirani kao izlazni podaci iz jednog ili više slojeva na strani kodera. Ti slojevi mogu biti od sloja 1 do sloja 5, koji odgovaraju brzinama prijenosa od 8 kbit/s do 32 kbit/s. To znači da je dekodeo uvjetovan brojem bitova koji se prime u svakom okviru. Sloj 1 (jezgreni sloj) te ACELP sloj (L2 sloj) se prvo dekodiraju i odrađuje se sinteza signala. Sintetizirani signal se provlači kroz pojasno propusni filter i ponovno uzorkuje frekvencijom od 16 kHz. Nakon prolaska kroz blok za oblikovanje šuma, takvom signalu se dodaje izlaz dobiven inverznom MDCT transformacijom. Zbrajanjem tih dvaju signala se dobiva novi signal nakon kojeg se primjenjuju novi inverzni težinski faktori kako bi se ponovno generirao širokopojasni signal koji je bio



originalno poslan. Konačno se primjenjuje filtriranje visokopropusnim filtrom. Filter iskorištava kašnjenje dekodera koji dolazi kao prirodna posljedica primjene MDCT-a, posebice u višim slojevima L3, L4 i L5. Filter na optimalan način kombinira dva signala, a to su signali dobiveni ponovnim uzorkovanjem (izlaz iz bloka za ponovno uzorkovanje) te signal dobiven primjenom inverznih težinskih faktora. Prvi signal je zapravo visokokvalitetni signal dobiven kao izlaz iz slojeva L1 i L2 iz podatkovnog toka, dok je drugi signal sintetički signal dobiven iz viših slojeva, a to su L3, L4 i L5. Ako je izlaz dekodera ograničen na slojeve L1, L2, i L3, pojasna širina se povećava na frekvencije u intervalu od 6400 Hz te 7000 Hz. Kada se dekodiraju slojevi L4 i L5, ne radi se povećanje pojasne širine budući da je tada u tom slučaju već cijeli spektar kvantiziran. U dekoderu se koristi posebna tehnika (eng. *advanced anti-swirling technique*) koja onemogućuje sintezu neprirodnih zvukova koji se događaju u pozadini, a prevladava im svojstvo stacionarnosti, tj. stalne prisutnosti tijekom govora. Primjer takvih zvukova je zvuk automobila. Ova tehnika smanjuje snagu i fluktuacije spektra pobudnog signala LP filtra. Neželjeni pozadinski zvukovi obično se događaju pri nižim brzinama prijenosa, te se stoga koristi i u nižim slojevima upravo iz tog razloga.

### 3.3.1 Princip rada dekodera

Funkcija dekodera se sastoji od dekodiranja primljenog ulaznog signala iz komunikacijskog kanala te njegove sinteze kako bi se dobio rekonstruirani govorni signal [4]. Tok signala će se detaljno obrađivati prema slici 2. Proces dekodiranja započinje s dekodiranjem sloja L1. ISF parametri se dekodiraju iz primljenog podatkovnog toka te se pretvaraju u ISP koeficijente (pomoću matematičke interpolacije) i posljedično u LP koeficijente. Ti koeficijenti se koriste za konstrukciju pobudnog signala koji će se koristiti zajedno s LP sintezom (filtriranje s LP sintetičkim filtrom) za rekonstruiranje govornog signala. Taj rekonstruirani signal se zatim podvrgava smanjenju neželjenih frekvencija (eng. *de-emphasis*), tj. procesu koji je inverzan procesu koji se nalazi na strani koda prije obrade govornog signala (eng. *pre-emphasis*). Pojačavaju se frekvencijski harmonici za koje se misli da pripadaju šumu. Tako se na taj način postiže bolji S/N omjer, tj. omjer signal/šum. Na kraju se primjenjuje pojačanje željenih harmonika u niskopojasnom području signala i signal se ponovno uzorkuje frekvencijom od 16 kHz. Tako se generira širokopojasni govorni signal koji ima frekvencijski spektar od 6400 – 7000 Hz. Taj signal se pridodaje ponovno uzorkovanom sintetičkom signalu kako bi se dobio željeni izlazni govorni signal. Prilikom procesa dekodiranja, postoji algoritamsko kašnjenje koje ovisi o frekvenciji uzorkovanja te broju dekodiranih slojeva [7]. Pri višim slojevima (L3-L5) se

iskorištava kašnjenje dekodera u iznosu od 10 ms kako bi se obavile operacije dodavanja i preklapanja pri inverznoj MDCT transformaciji. Pri nižim slojevima (L1 i L2) ne koristi se MDCT, a kašnjenje dekodera se može upotrijebiti za druge radnje kako bi se optimizirao izlazni govorni signal. Primjerice može se to vrijeme koristiti za rekonstrukciju nestalih okvira ili za poboljšanje kvalitete glazbe ukoliko se pored govora šalje istovremeno i signal koji u sebi sadrži glazbu i to pri brzinama od 8 do 12 kb/s. Algoritamsko kašnjenje za različite konfiguracije je dano u tablici 2, a podaci iz tablice su preuzeti iz literature [7].

Frekvencija uzorkovanja ulaznog signala	Frekvencija uzorkovanja izlaznog signala	Algoritamsko kašnjenje	
		Normalno dekodiranje	Dekodiranje pri niskim vrijednostima kašnjenja
16 kHz	16 kHz	42.875 ms	32.875 ms
8 kHz	8 kHz	43.875 ms	33.875 ms
16 kHz	8 kHz	42.8125 ms	32.8125 ms
8 kHz	16 kHz	43.9375 ms	33.9375 ms

**Tablica 1:** Algoritamsko kašnjenje za G.718 preporuku

Prema tablici 1, algoritamsko kašnjenje za slučaj kad se radi dekodiranje pri niskim vrijednostima kašnjenja, te vrijednosti su primjenjive za slojeve L1 i L2. Algoritamsko kašnjenje za slučaj normalnog dekodiranja vrijedi za slojeve L3, L4 i L5 u slučaju ako se žele koristiti prilikom kodiranja govornog signala.

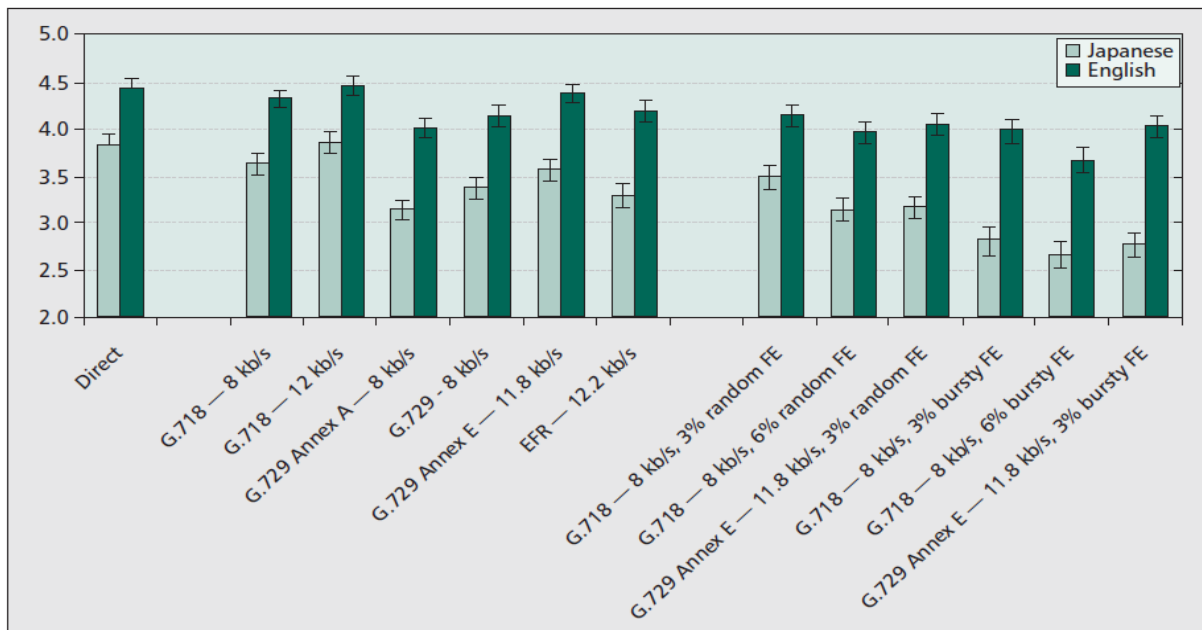
## 4. PRIMJENA I PREDNOSTI

Glavna primjena G.718 preporuke, odnosno kodeka, je u telefonskoj infrastrukturi [6]. Također se preporuka primjenjuje i u drugim područjima, poput video konferencije, video prijenosa te zvučnog (eng. *audio*) prijenosa. Glavna odlika preporuke je slojeviti dizajn. On omogućuje skalabilnost na način da se koriste slojevi koji su prikladni za neku određenu svrhu. Broj slojeva koji se koristi tijekom podatkovnog prijenosa ovisi o kapacitetu komunikacijskog kanala, trenutnih uvjeta prijenosa podataka, karakteristikama izvornog signala te dostupnim računalnim resursima prilikom procesiranja. Sljedeća prednost slojevitog dizajna kodeka je u mogućnosti iskorištavanja svojstva skalabilnosti kako bi se upravljalo zagušenjem na način da se šalju ili ispuštaju podaci koji pripadaju nekim od viših slojeva (tj. slojevima većima od L1). Također, slojeviti dizajn podržava mogućnost zaštite od grešaka na način da se koriste različite razine zaštite na jezgrenom sloju ili na višim slojevima. Što se tiče kontrole zagušenja mreže, G.718 preporuka kako je već rečeno, podržava kontrolu zagušenja na način da manipulira podatkovnim tokom (povećanjem ili smanjenjem podatkovnog toka). To se postiže u sprezi s RTP protokolom (eng. *Real-time Transport Protocol*). RTP protokol se koristi za slanje zvučnih i video poruka preko IP mreže [6]. Koristi se u komunikacijskim sustavima poput telefonije, videokonferencije itd. RTP okvir u sebi posjeduje zaglavlje i podatkovni segment koji se šalje mrežom (eng. *payload*). Pomoću RTP-a se postiže kontrola zagušenja mreže. RTP polje za podatak koji se šalje mrežom se iskorištava za tu kontrolu, a RTP nudi nekoliko različitih načina smanjenja pojasa (eng. *bandwidth*) kako bi se ostvarila što bolja kvaliteta kontrole zagušenja [6]. Najprikladniji mehanizam za kontrolu zagušenja je onaj koji ima najveći utjecaj na korisničko iskustvo, a ti mehanizmi ovise o strukturi polja za podatke u RTP okviru te o konfiguraciji sesije [6]. Načini kontrole zagušenja su zavisno od toga radi li kontrolu pošiljatelj ili neki čvor u mreži između pošiljatelja i primatelja. Prvi način je da RTP okviri koji sadrže podatke koji pripadaju višim slojevima se mogu ispustiti, zajedno sa svim povezanim transportnim blokovima. Drugi način je da transportni blokovi koji sadrže podatke najviših slojeva se mogu ispustiti. Što se tiče sigurnosti prijenosa podataka, su RTP paketi kojima se šalju podaci kodirani preporukom tj. kodekom G.718 kriptirani primjenom SRTP-a (eng. *Secure Real-time Transport Protocol*), koji nije ništa drugo nego oblik RTP protokola koji posjeduje dodatne značajke poput autentikacije poruka, povjerljivost te ponavljanje slanja okvira i konačno upravljanje zagušenja u mreži manipulacijom slanja te ispuštanja RTP paketa [6]. Budući da se kompresija primjenjuje *end-to-end*, nakon kompresije mora se primijeniti kriptiranje podataka. Također postoji opasnost

prijetnje od prisluškivanja podataka koji se šalju mrežom. Primjerice, napadač može ubrizgati (eng. *inject*) posebne pakete (eng. *datagrams*) u mrežu kojom se šalju podaci signala [6]. Ti paketi posljedično povećavaju opterećenje dekodera prilikom obrade podataka koji može postati preopterećen. Međutim, preporuka G.718 ima mehanizme zaštite za ovakav tip napada, budući da je većina tereta u smislu procesiranja podataka zapravo na strani koda. Međutim, ipak se preporuča korištenje SRTP-a u ovakvim situacijama.

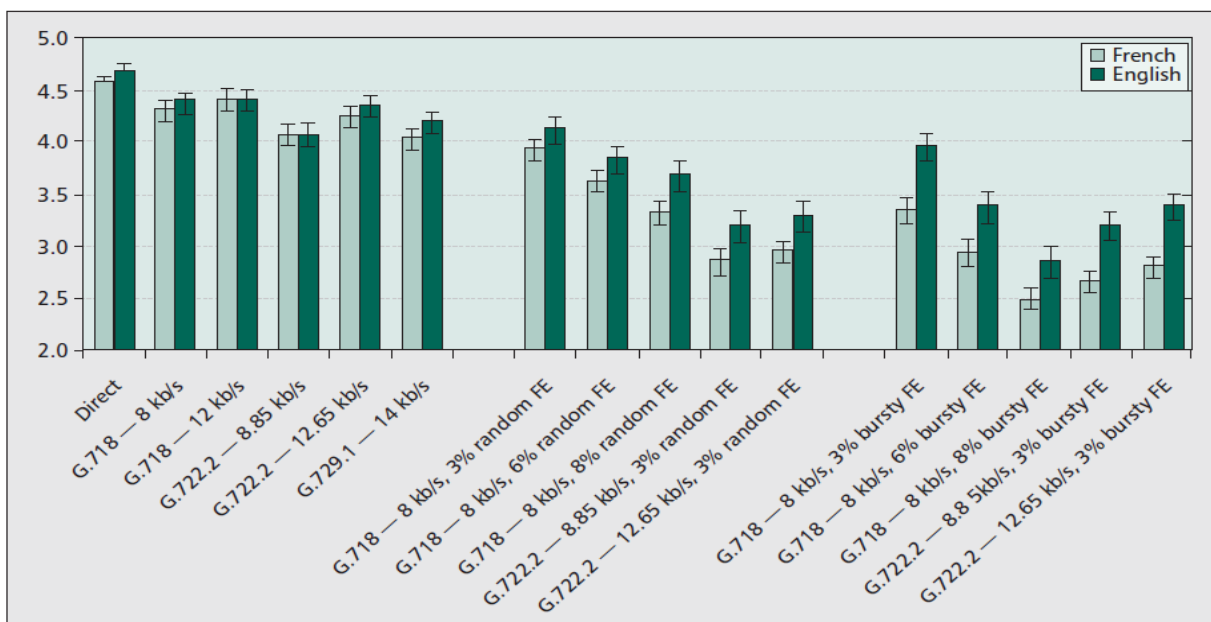
## 5. USPOREDBA S OSTALIM PREPORUKAMA

G.718 preporuka je formalno uspoređena s ostalim preporukama u smislu performansi rada 2009. godine, a testirao se širokopojasni govor, niskopojasni govor te glazbeni signali s obzirom na različite tipove ulaznog signala u kodek [7]. U mjerenju je sudjelovalo devet laboratorija koji su sudjelovali pri izradi preporuke, a uz evaluaciju su se koristili govorni signali na različitim jezicima. Postavljeno je 80 uvjeta od kojih je preporuka zadovoljila njih 78 u svim laboratorijima, a dva posljednja uvjeta je priznao samo jedan laboratorij. Testovi su pokazali da se najznačajniji napredak dogodio u području niskih brzina podatkovnog toka širokopojasnog govora. Također su pokazali da se značajan napredak dogodio i u području uskopojasnog govora. U potonjem slučaju, L1 sloj pri brzini od 8 kb/s je imao ponašanje slično G.729 preporuci koja je radila pri brzini od 11.8 kb/s pri kojoj se može čuti jasno govor u govornom signalu. Također, kodek je pokazao zadovoljavajuća svojstva u otežanim uvjetima (npr. buka) koja je bila prisutna unutar ulaznog signala kodeka za uskopojasni i širokopojasni govor. Testiranje se također provodilo i u slučaju ekstremnih uvjeta prijenosa gdje su izazvani gubici u prijenosu okvira te situacije gdje se dogodio veliki gubitak uzastopnih okvira u komunikacijskom kanalu [7]. Rezultati testiranja mogu se vidjeti za slučajeve prikazane na sljedećim slikama u usporedbi s ostalim preporukama, ispod čijih je dan opis usporedbe. Treba napomenuti da se za tumačenje rezultata koristila ACR metoda (eng. *Absolute Category Rating*) ITU-T preporuke P.800, a rezultati su prikazani koristeći MOS (eng. *mean opinion scores*). MOS je mjera za mjerenje kvalitete usluge u telekomunikacijskim sustavima, a predstavlja cjelokupnu kvalitetu sustava kroz aritmetičku sredinu svih vrijednosti ocjena koje su dali ispitivači koji su sudjelovali u testiranju govornog signala preporuke.



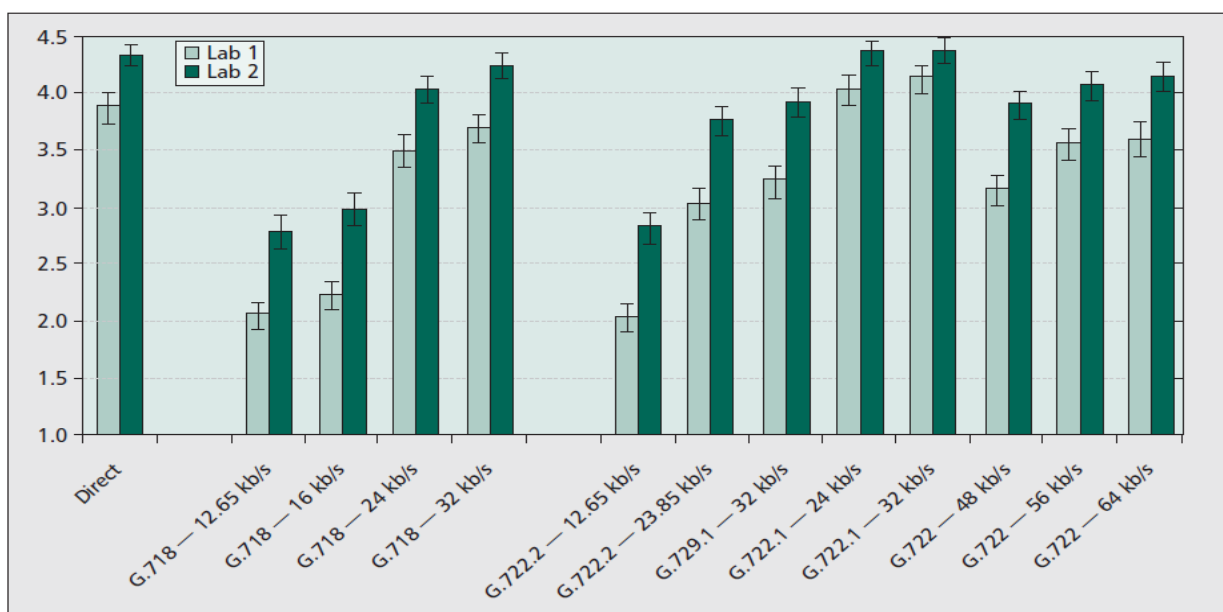
**Slika 3:** Subjektivni rezultati testova za uskopojasni govor

Slika 3 predstavlja MOS rezultate za uskopojasni govor kodiran pri brzinama od 8 kb/s i 12 kb/s u kanalu bez smetnji u usporedbi s ostalim preporukama. Na slici se također mogu vidjeti dobiveni rezultati testiranja s obzirom na vrstu govornog jezika koja se koristila za testiranje, a to su u ovom slučaju japanski jezik i engleski jezik. Za slučaj uskopojasnog govora može se reći da G.718 preporuka postiže rezultate usporedive ili čak i bolje u odnosu na neke od preporuka, ovisno koji slučajevi se promatraju.



**Slika 4:** Subjektivni rezultati za širokopojasni govor

Slika 4 predstavlja MOS rezultate za širokopolasni govor kodiran pri brzinama od 8 kb/s te 12 kb/s. G.718 preporuka, odnosno kodek je uspoređen s kodekom G.722.2 (pri brzinama kodiranja od 8.85 kb/s te 12.65 kb/s) te s G.729.1 (pri brzini od 14 kb/s). U uvjetima kanala bez smetnji G.718 kodek pri 8 kb/s pokazuje bolje performanse u odnosu na G.729.1 pri brzinama od 14 kb/s. Također se može promotriti iz slike da preporuka G.718 zadržava zadovoljavajuću kvalitetu govora čak uz stopu obrisanih okvira od 8 %. Konkretno G.718 pri brzini od 8 kb/s te s 8 % izgubljenih okvira ima bolje performanse od G.722.2 pri brzini od 12.65 kb/s u uvjetima gdje je postotak izgubljenih okvira za ovu preporuku bio na 3 %. Također kao i u prošlom slučaju može se vidjeti da su korišteni za testiranje japanski i engleski jezik.



**Slika 5:** Subjektivni rezultati za širokopolasni glazbeni signal

Slika 5 predstavlja rezultate širokopolasnog glazbenog signala za preporuku G.718 u usporedbi s ostalim preporukama. Rezultati koji su vidljivi za preporuku G.718 su rezultati dobiveni u režimu rada u kojem ta preporuka radi zajedno s preporukom G.722.2 pri brzini od 12.65 kb/s. Pokazano je također da se zadovoljavajuća kvaliteta glazbe može postići pri brzini kodiranja od najmanje 24 kb/s [7]. Za testiranje su korišteni rezultati MOS-a iz dvaju laboratorija koji su obavljali testiranje glazbenih signala.

## 6. ZAKLJUČAK

U ovom završnom radu predstavljena je G.718 preporuka gdje je opisan povijesni razvoj koji je doveo do razvitka ove preporuke, te je opisan princip rada i konačno je dana usporedba performansi te preporuke s rezultatima testiranja ostalih preporuka. Kako bi se unaprijedili telekomunikacijski sustavi, G.718 je opremljen opcijom da radi zajedno s G.722.2 preporukom pri brzini kodiranja od 12.65 kb/s. Rezultati testiranja preporuke G.718 su pokazali da su se postigle bolje performanse prvenstveno u kodiranju govornog i glazbenog signala u otežanim uvjetima gdje su prisutne smetnje u komunikacijskom kanalu. Trenutno se razvija dodatak koji će se također standardizirati te omogućiti preporuci G.718 reprodukciju zvučnih signala u frekvencijskom pojasu iznosa intervala od 50 – 14000 Hz. G.718 preporuka će se sigurno još dugo koristiti zbog svojih svojstava i performansi.



## LITERATURA

[1] <https://en.wikipedia.org/wiki/G.711>

[2] Codec, Wikipedia - <https://en.wikipedia.org/wiki/Codec>

[3] G.718, Wikipedia - <https://en.wikipedia.org/wiki/G.718>

[4] Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s, Recommendation ITU-T G.718, 2008.

[5] <https://en.wikipedia.org/wiki/Code>

[6] RTP Payload format for G.718 Speech/Audio draft-ietf-payload-rtp-g718-03 - <https://datatracker.ietf.org/doc/html/draft-ietf-payload-rtp-g718#section-4.1.2>

[7] G.718: A New Embedded Speech and Audio Coding Standard with High Resilience to Error-Prone Transmission Channels, Milan Jelinek i Tommy Vaillancourt, VoiceAge/University of Sherbrooke Jon Gibbs, Motorola, 2009.

[8] CELP, Wikipedia - [https://en.wikipedia.org/wiki/Code-excited\\_linear\\_prediction](https://en.wikipedia.org/wiki/Code-excited_linear_prediction)

[9] ACELP, Wikipedia - [https://en.wikipedia.org/wiki/Algebraic\\_code-excited\\_linear\\_prediction](https://en.wikipedia.org/wiki/Algebraic_code-excited_linear_prediction)

**Popis korištenih oznaka/kratice:**

ITU-T – International Telecommunication Union – Telecommunication Standardization Sector

CELP - Code-excited Linear Prediction

ACELP - Algebraic Code-excited Linear Prediction

VBR - Variable Bit Rate

CDMA - Code Division Multiple Access

EV-VBR - Embedded Variable Bit Rate

AMR-WB - Adaptive Multi-Rate Wideband

VMR-WB - Variable-Rate Multimode Wideband

MDCT - Modified Discrete Cosine Transform

FFT - Fast Fourier Transform

ISF - Immittance Spectral Frequencies

ISP - Immittance Spectral Pair

AbS - Analysis-by-Synthesis

LPC - Linear Predictive Coding

FEC - Frame Erasure Concealment

LP – Linear Prediction

HP - High Pass

RTP - Real-time Transport Protocol

SRTP - Secure Real-time Transport Protocol

MOS - Mean opinion scores

ACR - Absolute Category Rating

## SAŽETAK

Rad se bavi temom ITU-T preporuke G.718, koja je također i kodek. Koristi se u telekomunikacijskim sustavima za prijenos zvučnog signala IP mrežnu infrastrukturu. Prikazuje se povijesni razvoj koji je doveo do razvoja preporuke G.718. Zatim se prelazi na opis principa rada same preporuke gdje je dan detaljniji pregled funkcionalnosti te je napravljena podjela preporuke po bitnim cjelinama koje su dodatno obrađene. Kasnije je opisana primjena same preporuke i konačno je dana usporedba performansi G.718 preporuke zajedno s ostalim preporukama pri određenim uvjetima komunikacijskog kanala gdje je testiran zvučni signal, odnosno govorni i glazbeni signal.

### **Ključne riječi:**

G.718 preporuka, kodek, koder, dekodek, govorni signal, kompresija

## **ABSTRACT**

This thesis deals with ITU-T recommendation of G.718, which is also a codec. It is being used in telecommunications industry for transport of audio signal through IP network infrastructure. Firstly, a historical development has been shown which lead to development of G.718. Secondly the main principles of the recommendation have been given together with detailed overview of functionality divided into relevant units which have been documented accordingly. Thirdly the application of the recommendation has been given. Finally the performance overview has been given for the recommendation which are analysed together with other recommendations in various communication channel conditions for audio signals, primarily for speech and music signal.

### **Keywords:**

G.718 recommendation, codec, encoder, decoder, speech signal, compression

## ŽIVOTOPIS

Franko Kulešević rođen 27. svibnja 1997. godine u Zagrebu. Rano djetinjstvo proveo je u Izraelu, potom u Maleziji, gdje je razvijao svoje komunikacijske vještine i sposobnost prilagodbe multikulturalnom okruženju. Školovanje je započeo u osnovnoj školi „Franje Krežme“ u Osijeku. Pohađa pučko sveučilište gdje usavršava visoku razinu znanja engleskog jezika. Također, temeljni je korisnik njemačkog jezika. Upisuje i uspješno završava 1. (prvu) opću gimnaziju u Osijeku. Na osnovi odličnog uspjeha tijekom cijelog obrazovanja upisuje Fakultet elektrotehnike, računarstva i informacijskih tehnologija u Osijeku. Redoviti je student preddiplomskog stručnog studija računarstva. Također uspješno završava programerski tečaj za web dizajn na učilištu „Edunova“. Programski jezici kojima se služi i koje zna koristiti su osnove C, SQL, Java, HTML, CSS i baze podataka. Nogomet mu je najdraža slobodna aktivnost koja ga je pratila tijekom cijelog života i kroz koju je naučio timski rad i organizacijske vještine. Svjestan svoje ambicije upisao je PHP tečaj koji planira završiti kako bi bio bliži ispunjenju svoje karijere. Tijekom obrazovanja na fakultetu obavljao je praksu koja mu je pomogla pri profesionalnim i osobnom napretku. Volio bi raditi kao backend developer i programer.